

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 18:27:17 ; Search time 14.4545 Seconds
(without alignments)
107.116 Million cell updates/sec

Title: US-09-743-225-1
Perfect score: 30
Sequence: 1 LKTPRV 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phages.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 30 | 100.0 | 310 | 5 Q20160 | Q20160 caenorhabdi |
| 2 | 30 | 100.0 | 405 | 4 Q36E86 | Q36E86 homo sapien |
| 3 | 30 | 100.0 | 406 | 16 Q8EKZ4 | Q8EKZ4 oceanobacil |
| 4 | 30 | 100.0 | 435 | 11 Q8VCA5 | Q8VCA5 mus musculu |
| 5 | 30 | 100.0 | 939 | 4 Q15071 | Q15071 homo sapien |
| 6 | 30 | 100.0 | 949 | 4 Q60369 | Q60369 homo sapien |
| 7 | 30 | 100.0 | 988 | 4 Q8WUF7 | Q8WUF7 homo sapien |
| 8 | 30 | 100.0 | 1067 | 11 Q8CH09 | Q8CH09 mus musculu |
| 9 | 30 | 100.0 | 1067 | 11 Q8BY32 | Q8BY32 mus musculu |
| 10 | 30 | 100.0 | 1082 | 4 Q8IX01 | Q8IX01 homo sapien |
| 11 | 29 | 96.7 | 38 | 16 Q9KGN8 | Q9KGN8 vibrio chol |
| 12 | 29 | 96.7 | 147 | 15 Q8XJ82 | Q8XJ82 clostridium |
| 13 | 29 | 96.7 | 1383 | 10 Q9LVH0 | Q9LVH0 arabidopsis |
| 14 | 28 | 93.3 | 174 | 11 Q64683 | Q64683 mesocricetu |
| 15 | 28 | 93.3 | 263 | 17 Q8ZTE5 | Q8ZTE5 pyrobaculum |
| 16 | 28 | 93.3 | 468 | 4 Q9BX59 | Q9BX59 homo sapien |

| | | | | | |
|----|----|------|-----|-----------|---------------------|
| 17 | 28 | 93.3 | 468 | 4 Q9NWB8 | Q9NWB8 homo sapien |
| 18 | 28 | 93.3 | 549 | 4 Q8WW11 | Q8WW11 homo sapien |
| 19 | 28 | 93.3 | 549 | 4 Q9NX50 | Q9NX50 homo sapien |
| 20 | 28 | 93.3 | 549 | 4 Q96QG7 | Q96QG7 homo sapien |
| 21 | 28 | 93.3 | 775 | 13 Q91065 | Q91065 lophius ame |
| 22 | 27 | 90.0 | 41 | 16 Q8EJ26 | Q8EJ26 shewanella |
| 23 | 27 | 90.0 | 109 | 16 Q8D251 | Q8D251 wiggleswort |
| 24 | 27 | 90.0 | 144 | 16 Q8XTG1 | Q8XTG1 raistonia s |
| 25 | 27 | 90.0 | 177 | 16 Q9CHP0 | Q9CHP0 lactococcus |
| 26 | 27 | 90.0 | 245 | 5 Q19964 | Q19964 caenorhabdi |
| 27 | 27 | 90.0 | 277 | 2 Q8VTR6 | Q8VTR6 hafiaia alve |
| 28 | 27 | 90.0 | 233 | 5 Q9W465 | Q9W465 drosophila |
| 29 | 27 | 90.0 | 363 | 2 Q8VP95 | Q8VP95 burkholderi |
| 30 | 27 | 90.0 | 363 | 9 Q8W6S0 | Q8W6S0 bacterioph |
| 31 | 27 | 90.0 | 382 | 2 Q93M46 | Q93M46 legionella |
| 32 | 27 | 90.0 | 388 | 16 Q99SX8 | Q99SX8 staphylococ |
| 33 | 27 | 90.0 | 388 | 16 Q8NV59 | Q8NV59 staphylococ |
| 34 | 27 | 90.0 | 392 | 2 Q8RNB5 | Q8RNB5 bartonella |
| 35 | 27 | 90.0 | 392 | 16 Q9F3M6 | Q9F3M6 streptomyce |
| 36 | 27 | 90.0 | 433 | 5 Q965T3 | Q965T3 caenorhabdi |
| 37 | 27 | 90.0 | 435 | 5 Q9Y016 | Q9Y016 paracentrot |
| 38 | 27 | 90.0 | 492 | 5 Q45457 | Q45457 caenorhabdi |
| 39 | 27 | 90.0 | 525 | 5 Q95861 | Q95861 drosophila |
| 40 | 27 | 90.0 | 525 | 5 Q9V340 | Q9V340 drosophila |
| 41 | 27 | 90.0 | 556 | 5 Q9VFG2 | Q9VFG2 drosophila |
| 42 | 27 | 90.0 | 568 | 16 Q8K6R6 | Q8K6R6 streptococc |
| 43 | 27 | 90.0 | 604 | 4 Q43606 | Q43606 homo sapien |
| 44 | 27 | 90.0 | 613 | 16 Q97RV5 | Q97RV5 streptococc |
| 45 | 27 | 90.0 | 613 | 16 Q99YV1 | Q99YV1 streptococc |

ALIGNMENTS

RESULT 1

Q20160 ID Q20160 PRELIMINARY; PRT; 310 AA.
AC Q20160;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 35.6 kDa protein.
GN F38E1.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Gattung S., Le T.T.;
RT "The sequence of C. elegans cosmid F38E1.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN EMBL; U41996; AAA83477.1;
DR HSSP; Q06486; ICKI
DR WormPep; F38E1.3; CE04522.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 KW PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Hypothetical protein; ATP-binding; Transferrase.
 SQ SEQUENCE 310 AA; 35553 MW; 98C00C832A507AC1 CRC64;

Query Match 100.0%; Score 30; DB 5; Length 310;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
 |||||
 Db 185 LKTPRV 190

RESULT 2
 ID Q96E86 PRELIMINARY; PRT; 405 AA.
 AC Q96E86;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to transmembrane protease, serine 4 (Fragment).
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pancreas;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; BC012752; AAH12752.1; -;
 DR HSP; P00761; IAN1.
 DR MEROPS; S01.034; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLA; 1.
 DR SMART; SM00202; SR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolyase; Protease; Serine protease; Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 405 AA; 44474 MW; 951ACD52D9D48E04 CRC64;

Query Match 100.0%; Score 30; DB 4; Length 405;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
 |||||
 Db 168 LKTPRV 173

RESULT 3
 ID Q8EK24 PRELIMINARY; PRT; 406 AA.
 AC Q8EK24;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hemolysin.
 GN OB3438
 OS Oceanobacillus iheyensis.
 OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.

Query Match 100.0%; Score 30; DB 4; Length 405;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OX NCBI_TaxID=182710;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE-22220767; PubMed-12235376;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; AF004604; BAC13394.1; -;
 KW Complete proteome.
 SQ SEQUENCE 406 AA; 45787 MW; 00BD195D5670D53C CRC64;

Query Match 100.0%; Score 30; DB 16; Length 406;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
 |||||
 Db 198 LKTPRV 203

RESULT 4
 Q8VCA5 PRELIMINARY; PRT; 435 AA.
 ID Q8VCA5;
 AC Q8VCA5;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to transmembrane protease, serine 4 (Channel-activating
 DE protease 2).
 GN Tmprs54.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast tumor;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-22144321; PubMed-12149280;
 RA Vuagniaux G., Vallet V., Jaeger N.F., Hummler E., Rossier B.C.;
 RT "Synaptic Activation of ENAC by Three Membrane-bound Channel-
 RT activating Serine Proteases (mCAP1, mCAP2, and mCAP3) and Serum- and
 RT Glucocorticoid-regulated Kinase (Sgk1) in Xenopus Oocytes";
 RL J. Gen. Physiol. 120:191-201(2002).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; BC021368; AAH21368.1; -;
 DR EMBL; AY043240; AAK85307.1; -;
 DR HSP; P00761; IAN1.
 DR MEROPS; S01.034; -;
 DR MGD; MGI:2384877; Tmprs54.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolyase; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 435 AA; 47495 MW; DC52E45A43E01369 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 435;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 LKTPRV 6
DB 198 LKTPRV 203

RESULT 5
O15071 PRELIMINARY; PRT; 939 AA.
AC O15071;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA0365 (Fragment).
GN KIAA0365.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=9734984; PubMed=9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
DR EMBL; AB002363; BAA20820.1; -.
DR InterPro; IPR000467; G_patch.
DR InterPro; IPR000061; Surp.
DR Pfam; PF01585; G_patch; 1.
DR Pfam; PF01805; Surp; 2.
DR SMART; SM00443; G_patch; 1.
DR SMART; SM00648; SWAP; 2.
DR PROSITE; PS0174; G_PATCH; 1.
DR Hypothetical protein.
KW NON_TER 1
SQ SEQUENCE 939 AA; 103678 MW; B31CA2E3A9D4EBDC CRC64;

Query Match 100.0%; Score 30; DB 4; Length 939;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
DB 250 LKTPRV 255

RESULT 6
O60369 PRELIMINARY; PRT; 949 AA.
AC O60369;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA0365 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Ganes J., Danganan L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA Carrano A.V.;
RT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in

```

```

19p12.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004447; AAC06129.1; -.
DR InterPro; IPR000467; G_patch.
DR InterPro; IPR000061; Surp.
DR Pfam; PF01585; G_patch; 1.
DR Pfam; PF01805; Surp; 2.
DR SMART; SM00443; G_patch; 1.
DR SMART; SM00648; SWAP; 2.
DR PROSITE; PS0174; G_PATCH; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 949 AA; 105111 MW; D64AA847DECB6F5A CRC64;

Query Match 100.0%; Score 30; DB 4; Length 949;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
DB 250 LKTPRV 255

RESULT 7
O8WUF7 PRELIMINARY; PRT; 988 AA.
AC O8WUF7;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Duoenum;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020586; AAH20586.1; -.
DR InterPro; IPR000467; G_patch.
DR InterPro; IPR000061; Surp.
DR Pfam; PF01585; G_patch; 1.
DR Pfam; PF01805; Surp; 2.
DR SMART; SM00443; G_patch; 1.
DR SMART; SM00648; SWAP; 2.
DR PROSITE; PS0174; G_PATCH; 1.
KW Hypothetical protein.
SQ SEQUENCE 988 AA; 110412 MW; 8C5B29A88BD77449 CRC64;

Query Match 100.0%; Score 30; DB 4; Length 988;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKTPRV 6
DB 341 LKTPRV 346

RESULT 8
O8CH09 PRELIMINARY; PRT; 1067 AA.
AC O8CH09;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Arginine/serine-rich 14 splicing factor.
GN SFRS14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-C57BL/6J;
RA  Sampson N.D., Hewitt J.E.;
RT  "Cloning of the mouse splicing factor, SFRS14.";
RL  Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF518875; AAN77118.1; -
SQ  SEQUENCE 1067 AA; 118116 MW; BEA25F1CE71C4D92 CRC64;

Query Match      100.0%; Score 30; DB 11; Length 1067;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 LKTPRV 6
Db  383 LKTPRV 388

RESULT 9
Q8BY32 PRELIMINARY; PRT; 1067 AA.
ID AC Q8BY32;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical D111/G-patch domain/aminocyl-transfer RNA synthetases
DE class-I/glutamic acid-rich region/swap / SURP containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK042293; BAC31218.1; -
KW Hypothetical protein.
SQ SEQUENCE 1067 AA; 118106 MW; CE191DC7C31C4A55 CRC64;

Query Match      100.0%; Score 30; DB 11; Length 1067;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 LKTPRV 6
Db  383 LKTPRV 388

RESULT 10
Q8IX01 PRELIMINARY; PRT; 1082 AA.
ID AC Q8IX01;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Arginine/serine-rich 14 splicing factor.
GN SFRS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Sampson N.D., Hewitt J.E.;
RT "Cloning of the novel splicing factor, SFRS14.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF518874; AAN77117.1; -
SQ SEQUENCE 1082 AA; 120237 MW; BFDCB6EF096FA736 CRC64;

```

```

Query Match      100.0%; Score 30; DB 4; Length 1082;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 LKTPRV 6
Db  393 LKTPRV 398

RESULT 11
Q9KSN8 PRELIMINARY; PRT; 38 AA.
ID AC Q9KSN8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein VCL218.
GN VCL218.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004202; AAF94377.1; -
DR TIGR; VCL218; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 38 AA; 4085 MW; F5786335FC4E6D98 CRC64;

Query Match      96.7%; Score 29; DB 16; Length 38;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 LKTPRV 6
Db  14 LKTPRI 19

RESULT 12
Q8XJ82 PRELIMINARY; PRT; 147 AA.
ID AC Q8XJ82;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein CPE1879.
GN CPE1879.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogawara N., Hattori M., Kubara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002);

```

```
DR EMBL; AF003192; BAB81585.1; -.
DR Pfam; PF04463; DUF523; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 147 AA; 15718 MW; 629908F16F8156C0 CRC64;

Query Match          96.7%; Score 29; DB 16; Length 147;
Best Local Similarity 83.3%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKTPRV 6
DB 45 LKTPRI 50

RESULT 13
Q9LVHO PRELIMINARY; PRT; 1383 AA.
AC Q9LVHO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DNA-directed RNA polymerase II largest chain.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20181125; PubMed=10718137;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and PAC
RT clones."
RL DNA Res. 7:31-63(2000).
RL EMBL; AB019231; BAA96933.1; -.
DR InterPro; IPR006592; RNA_pol_A.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOLA.N; 1.
KW DNA-directed RNA polymerase.
SQ SEQUENCE 1383 AA; 153939 MW; 9D091923B6A1BCC8 CRC64;

Query Match          96.7%; Score 29; DB 10; Length 1383;
Best Local Similarity 83.3%; Pred. No. 6,7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKTPRV 6
DB 1152 LKTPRI 1157

RESULT 14
Q64683 PRELIMINARY; PRT; 174 AA.
AC Q64683;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CAD protein carbamylphosphate synthetase domain (Fragment).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
```

```
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
SQ SEQUENCE FROM N.A.
RX MEDLINE=93068573; PubMed=1359654;
RA Musmanno L.A., Jamison R.S., Barnett R.S., Buford E., Davidson J.N.;
RT "Complete hamster CAD protein and the carbamylphosphate synthetase
RT domain of CAD complement mammalian cell mutants defective in de novo
RT pyrimidine biosynthesis."
RL Somat. Cell Mol. Genet. 18:309-318(1992).
DR EMBL; S48735; AAB24160.2; -.
DR HSSP; P00907; IA9X.
DR InterPro; IPR002474; CPSase_sm_chain.
DR Pfam; PF00988; CPSase_sm_chain; 1.
FT NON_TER 174 174
SQ SEQUENCE 174 AA; 18606 MW; 3626118E21264655 CRC64;

Query Match          93.3%; Score 28; DB 11; Length 174;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKTPRV 6
DB 164 IKTPRV 169

RESULT 15
O8ZTE5 PRELIMINARY; PRT; 263 AA.
ID O8ZTE5
AC O8ZTE5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polyprenyl synthetase.
GN PAE3294.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009922; AAL64817.1; -.
DR InterPro; IPR000092; Polyprenyl_synt.
DR Pfam; PF00348; polyprenyl_synt; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Complete proteome.
SQ SEQUENCE 263 AA; 29616 MW; AE265035F542FF67 CRC64;

Query Match          93.3%; Score 28; DB 17; Length 263;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKTPRV 6
DB 98 IKTPRV 103

Search completed: August 28, 2003, 18:37:52
Job time : 16.4545 secs
```